

Comments to the 1st PRD draft on W+jets, CDF 10937  
Jonathan Rosner, Diego Tonelli and Giorgiob, November 16, 2012  
(with reference to November 2 one column version with line numbers)

### GENERAL COMMENTS

We find that the paper fails to dig deep into the core motivations for this measurement and the details of the measurement approach, which are important to reproduce it and to compare it with theory and existing measurements. In general, it seems too poor of details and generic for such an important topic.

In the introduction the authors should explain why this final measurement by CDF is important and what it does add to the current experimental and phenomenological landscape of this physics. An expanded discussion should add details on:

- what the muon channel adds to the electrons one (not just statistics, but different/lower backgrounds etc.)
- what this measurements adds to the previous CDF measurement (besides stat.).
- why is this measurement at CDF important while there are much more precise measurements recently released by ATLAS and CMS (pp-vs ppbar? Different systematics? Bottom-line: what do we gain?)

In the final discussion one should compare these results with previous CDF and D0 measurements and with the current ATLAS and CMS measurements. The eight - lines-only conclusion is clearly insufficient.

Descriptions of many crucial steps in this analysis are just overviewed. They need to be expanded with quantitative details that allow the reader to form an opinion on the solidity of the measurement and on how it compares with the rest of the world. This includes

- expanding the discussion of the unfolding procedure (see detailed comments);
- adding ex-novo a discussion on how exactly the MC programs used to predict theory were configured and tuned in this analysis (renormalization/hadronization scales? cuts in Pythia-Alpgen matching?);
- adding ex-novo a discussion on the theory uncertainties (related to the choices above) and show them in the plots;
- adding plots that show the sample composition in the signal and control samples;
- adding plots that show the ratio between our data and the theory normalized to the uncertainty, so that we see to how many  $\sigma$ 's the discrepancies amount to;
- In general, a more detailed discussion of each plot describing the results. They are just shown with no discussion on what we learn from them and why.

The paper is well written but the text pays little attention to finding the appropriate expressions for the message being conveyed. Here we add a number of general suggestions, while detailed ones are added line-by-line.

Figures quality is poor. We understand that these are place-holders for figures of higher quality. The figures' captions need in general to be expanded to describe exactly what the figure displays, including a description of the various line styles, color codes etc. Can look at past W+nj paper for examples. It's clear that the captions of Fig 1-4 are uninformative.

Suggest using  $c$  (not  $c$ ) for the speed of light (everywhere).

Capitalize only the first word in Tables' headings and row's labels (see also detailed comments). Do not capitalize "quantum chromodynamics"

Pay attention to the conceptual distinction between a "charged particle" which has  $p_t$ , charge, etc., and its trajectory, the "track", which has curvature,  $\phi$ , etc. Often the two things are mixed up. We spotted a few of them, but a systematic check may be beneficial.

The notation for the differential cross section for "n" jets that you use in the final tables may be confusing. Consider  $(d\sigma/dp_t)_n$ .

In the Tables of results (Tab IX-XX), uncertainties may have strongly different orders of magnitudes, e.g. for the first entry of Tab IX you write  $1520 \pm 1.2 +79 -76 \pm 91$ . The significant digits of these numbers should be rounded according to the standard rules:  $1520 \pm 1 +79 -76 \pm 90$ . There are many of these cases.

In Tables XIII-XX,  $p_t$  ranges (e.g., 20-25, 25-30,...) are expressed with a single minus sign, while you should use two of them  $\rightarrow$  20--25, 25--30...

References have several issues. First and middle names of authors need to be abbreviated. Standard format should be used for journal abbreviations as per SPRG guide <http://gate.hep.anl.gov/abw/reader/check.pdf>. For example [2] should be "J. of High Energy Physics 01 (2009) 012."

Try to avoid, or limit, the use of "our" in a physics paper.

Try to avoid "cut" in favor of "selection requirement/criterion" whenever possible.

Try to avoid "error" in favor of "uncertainty" whenever possible

Suggest using " $B$ " instead of "Br" for branching ratio symbol.

#### LINE BY LINE

#### Abstract

**Lines 2 and 3.** Suggest: "...in association with a W boson, identified through its decay to an electron or a muon and a neutrino. The transverse momentum spectra of the jets are also reported."

**Line 4.** "...and correspond to an integrated luminosity of..."

**Line 5 and 6.** Suggest: "The measurements are in good agreement with predictions based on perturbative Quantum Chromodynamics as implemented in the ALPGEN+PYTHIA and BLACKHAT+SHERPA simulation programs.",

**Lines 7 to 9.** Suggest moving this sentence to the second line, after "...in association with the W boson."

## Page 8.

**Line 3.** "...1.96 TeV, and (NUMBER)% of them decay..."

**Line 4.** To avoid an unpleasant repetition, suggest "can also be produced in the same events" rather than "can be produced in association with the W boson".

**Line 7.** Suggest "allow a stringent test of these predictions".

**Line 8.** "...incisively"? The motivations should be expanded into a more articulated discussion. The very important role of this measurement for gauging the backgrounds in a broad set of high-pt searches for new physics and precision measurements of SM processes like top and Higgs should be mentioned. Suggest "theoretical calculations" rather than "theoretical predictions".

**Line 9.** Suggest "... (NLO), and results can be obtained..."

**Line 17.** Suggest "...we reduce the rate uncertainties by normalizing..."

**Line 19.** Suggest "...on the theoretical predictions."

**Lines 20, 21.** Suggest "...were obtained from data collected... (CDF) with an integrated collider luminosity of  $2.8 \text{ fb}^{-1}$ , with..."

**Line 21.** It is not clear what "with both electron and muon channels" refers here since you haven't mentioned the W decay channels yet.

**Line 23.** Use "D0" (their official name)

**Line 25.** Suggest "from ours." rather than "from the CDF analysis."

**Line 26, 27.** Remove extra space before [10]. Suggest "...[10] in several ways.". Also, delete "for the first time in CDF" (unnecessary, and the address to "the muon channel" sounds generic). "Second,...,?" . The difficulties due to higher luminosities are not an improvement, reconstruction of the event vertex is, as a necessary reaction to a limitation of the experimental technique we use. This also holds for the last point you make (Page 9, line 1) on this analysis having more backgrounds than  $Z \rightarrow \ell\ell$ . Please rearrange the sentence.

This discussion that starts at line 26 should be expanded as per general comments. Please motivate why we added muons (not only for more statistics), and what this measurements adds to what was done in the previous CDF measurement. Also here you should mention the recent measurements from CMS and ATLAS, and discuss what we learn from this measurement with respect to what they did.

**Line 28.** Suggest "...collisions occur during..."

**Line 30.** Suggest "...to minimize this contamination and we..."

**Line 31, and first lines of page 9.** Suggest: "Backgrounds contributed by multijets and top quark pair events are not negligible in  $W \rightarrow \ell\nu$  processes as they are in  $Z \rightarrow \ell\ell$  events [12]. Accordingly, we developed..."

## Page 9

**Lines 4 to 11.** Please use "Sec." rather than "Section", except when it is the first word following a period.

**Line 11.** Suggest removing the first "final".

**Line 12.** "The CDF II detector"

**Line 15.** "...charged-particle tracking" (hyphens). Suggest "...surrounded by projective..."

**Line 23.** "...we use the polar angle  $\theta$  and the azimuthal angle  $\phi$ ."

**Line 25.** "is defined as"

**Line 25.** Suggest "...the distance between reconstructed physics objects in the..."

**Lines 26 and 27.** "... where  $\Delta\eta$  and  $\Delta\phi$  are the differences..."

## Page 10

**Line 2.** Suggest "...consists of a silicon microstrip detector [15] mounted..."

**Line 4.** Suggest "... arranged in a barrel geometry..."

**Lines 6, 7.** Discussing signal-to-noise in this context might be too technical. Consider removing "that has good...doses."

**Line 13.** You have "resolution" twice. Suggest using as customary  $\sigma(p_t)/p_t^2$ .

**Line 23.** Suggest "...are installed in the electromagnetic calorimeter near the..."

Please check the PHA resolution. I thought that the statistical term is worse than the CHA, around  $80\%/\sqrt{E}$ .

**Line 24.** "...the maximum development of an electromagnetic shower..."

**Lines 27 and 28.** Suggest "The outer drift chambers ...outside the 60 cm thick magnet steel yoke."

## Page 11

**Line 6.** "that is important" is unnecessary

**Line 8.** "cross sections"

**Line 9.** "Cherenkov counters"

**Line 11.** Suggest "On-line event selection system"

**Line 12.** "The on-line event selection (trigger) and data acquisition system is..."

**Line 15.** "events during"

**Line 16.** Quoting 132 ns here may be misleading. It's true that the pipeline was designed to be synchronous at 132 ns, but the actual bunch-to-bunch time distance has always been at 396 ns. Please fix this somehow.

**Line 17.** Suggest "combined" rather than "dual"

**Line 21.** Suggest ...tens of Hertz. Events accepted by the trigger are written..."

**Line 22.** Suggest "...reconstruct "objects" like..., to be used in the analysis of physical processes."

**Line 24.** It reads awkward. Suggest "We begin with a 2.8 fb<sup>-1</sup> data sample selected on-line by..."

**Line 25.** There is no need to specify again the detectors in parenthesis. Suggest deleting.

**Line 26.** "...searching for W bosons that..."

## Page 12.

**Line 4.** Suggest "Event vertex requirements" (no capitals except on the first word. The same in all headings till the end of page 15).

**Line 5.** “We bound the maximum number...”

**Line 6.** “...may produce jets not associated with the hard interaction that produced the  $W$  boson”.

**Line 7.** “...few low- $p_t$  particles are irrelevant...”

**Line 8.** “...all charged particles associated with...”

**Line 11.** “...only if at least one and no more than three good-quality collision vertices are reconstructed.”

**Line 12.** You should mention the criterion adopted for choosing the event vertex.

**Line 13.** “...charged-lepton track...”

**Lines 18 and 19.** Suggest “...we expect a charged track pointing to an energy cluster in the electromagnetic calorimeter, where most of the electron energy should be deposited.”

**Line 19.** “We search for ...”. What is “a single electron”? You are trying to define an electron, but you only have tracks available. Please rephrase as “To select a candidate electron sample, we...”

**Line 20.** You look for a track directed to the fiducial region of the electromagnetic calorimeter, and set a lower limit to the track momentum. Suggest rephrasing as “...electron whose trajectory points to the fiducial acceptance region of the calorimeter...”.

**Line 21.** “electron candidates”.

**Line 23.** Suggest “...not to exceed 0.055...”.

**Line 24.** “...the energy of the electromagnetic cluster...”

**Line 28.** Delete both “Our” and just say “The selection...mirrors the electron-channel selection requirements”.

## Page 13

**Line 1.** Suggest “We search for a track that geometrically matches a track segment in the muon detectors and is associated to an energy deposit in the calorimeter consistent with....”

**Line 2.** “...and require that the energy deposited in the calorimeter be consistent with that expected for...”

**Line 3.** “...single muon whose trajectory points to the fiducial acceptance region of the central muon...”

**Line 7.** “depositions”

**Line 8.** The “impact parameter” was not defined. “ $d_0$ ” does not define what it is.

**Table I.** “Event vertex selection requirements”, “Selection requirements”, Notes and details” (no capitals after first word). Also “electron” and “muon” in Tables II and III. See general comments for capitalization in the headings and row labels. You could also consider merging the three tables together into a single “Event selection” table to make this more compact.

## Page 14

**Line 3.** “and other seed-based cone algorithms” is unnecessary and can be deleted. Also “infrared- and collinear-safe”.

**Line 4.** “to all orders in  $\alpha_s$  in pQCD”, and “is an improvement” over what? Please add details.

**Line 5.** Suggest “...center of seed clusters which in turn provide a better treatment of...”

**Line 7.** Suggest “ region in which the jet-energy resolution is best understood...”

**Line 13.** Suggest adding "(hadron level corrections [26])" after "accurately" and deleting the following sentence.

**Line 14.** “We calculate...”. This discussion of the unfolding seems too brief given the relevance this procedure has for the final results. Please add details on

- how the ALPGEN+Pythia samples are generated (which configurations? Which scales? Which pdf? Which matching cuts?)

- how do you define long lived hadrons (which tau?). Are  $\mu/e/\nu$  from B decays included in them or not?

- how do you deal with dependences of the unfolding constants on the true MC spectrum used? Did you do any iterative procedure to converge to corrections that are universal enough to cover differences between the true MC and the true data spectrum?

Also, description of the unfolding would be more visible and easier to follow if moved in Section V (cross-section definition).

**Line 21.** Suggest ‘is found to increase approximately logarithmically’

## Page 15

**Line 4.** Should one use the singular: “from any jet”?

**Line 6.** Reducing contaminations from fake *jets*? Please revise.

**Line 14.** It's not that the "imbalance is best measured in the transverse plane". It's just that this is the only plane where you can define an imbalance, using the conservation of total  $p_t = 0$ .

**Line 19.** “...is corrected for energy scale biases and any...”

**Line 21.** “ No direct selection requirement is imposed on  $E_{t,miss}$ .”

**Line 22.** Suggest putting the formula for transverse mass here. It's an important variable in this work and it takes just half a line.

## Page 16.

**Line 1.** “while retaining most of the signal” should go to the end of sentence. Also “...below the known  $W$  mass”.

**Line 3.** Suggest quoting the muon limit first.

**Line 5.** “Three classes of backgrounds contribute:...”

**Line 13.** “...backgrounds. The other component...”

**Line 14.** “...such a way that only one parameters is determined by the fit, the fraction of...”. Please add details on how the uncertainties from the subtraction are accounted for in the template fit.

**Line 22.** “Top pair and diboson background estimation” (no capitals. The same in the titles of the other paragraphs of the chapter.

**Line 23** “were estimated”

## Page 17

**Tab IV.** “Background cross section.”. See general comments.

**Line 3.** Suggest “The process  $Z(\rightarrow l^+ l^-) + \text{jets}$  contributes...”

**Line 5.** Delete "on the other hand"

**Line 6.** No comma after "formed", and add comma after " $Z \rightarrow l^+ l^-$  background"

**Line 9.** Suggest "by construction"

**Line 12.** "...two of the standard..."

**Line 13.** "...which may contribute...", "...taken into account by using MC estimates..."

**Line 14.** No quotes on "contamination"

**Line 16.** In section C please add all plots of the  $M_T$  fits with sample composition or at least some examples of them for the  $\mu$  and  $e$  channels.

## Page 18

**Tables V and VI.** Suggest "Estimated background fractions...", and please add all uncertainties.

**Table VI.** "...as in Table V..."

**On line preceding Eq.1:** "...in Sec. III C,..." (add space)

**On 5<sup>th</sup> line above the equation,** comma after "e.g."

**Line above Eq. (1):** "as follows:" (colon)

## Page 19

**Line 2.** Suggest deleting "requirements"

**Lines 3 and 4.** Suggest "...from the kinematics of the simulated signal and the geometry of the detector, after imposing the generator-level cut listed in table VII."

**Line 7.** Suggest "...cross sections which makes them more easily comparable to theoretical predictions..." A detailed discussion of where the efficiency ( $\epsilon$ ) is extracted from is missing. Consider adding it here or somewhere else.

**Line 14.** "... (4,2%), and in the total... Adding these uncertainties in quadrature yields a total uncertainty..."

**Line 20.** "... induce slight changes..."

## Page 20

**Table VIII row headings:** capitalize only first word in each row (see general comments).

**Line 7.** "simulations"

**Line 12.** "...weights to simulated events to increase the agreement."

## Page 21.

When you discuss the  $t$ - $\bar{t}$  cross section, please show that the rate you subtract from your data matches the CDF measurements.

Please discuss here the systematic uncertainties due to the unfolding, in particular the dependence on the true spectrum of the unfolding parameters.

**Line 27.** "<sup>th</sup>"

**Line 29.** See general comments. You should provide here way more details on the specific configurations you used to produce these ALPGEN+PYTHIA and BLACKHAT+SHERPA samples and why you chose them and not something else (I am referring to choice of PDF, renormalization/hadronization scales, matching cuts etc..) Also, you should discuss on the uncertainties associated to the predicted cross sections, which are related to the arbitrariness of the above choices.

## Page 22.

**Fig. 1.** You should provide the corresponding ratio plot between data and theory. As it is one cannot understand if there is agreement of discrepancy in the first bin. If there's a discrepancy it could equally be  $3\sigma$  as well as  $300\sigma$ .

**Caption.** Add "events" at the end of the current caption. However, this caption should be greatly expanded explaining what enters in each plot and the various lanes and color codes. Basically you should report and explain what's in the plot's label.

**Line 2.** "is denoted by"

**Line 6.** Please comment on the agreement/disagreement of the two MC with the data. In what they differ? What the disagreement could be due to? What do we learn from this?

**Line 9.** "B. Jet multiplicity ratio" (no capitals). Same in C below.

**Line 14.** Same comment as on line 6 above.

## Page 23

**Fig 2.** Same comments as for Fig 1.

**Line 4.** "...n<sup>th</sup> leading jet..."

**Line 5.** "...for  $n \geq 3$ , or  $\geq 4$ ..."

**Line 6.** "estimations"

## Page 24.

**Fig 3.** Add plots showing the data/theory ratio. It's impossible to form an opinion on whether there is agreement or not and by how much, at leads for the first two plots.

Expand caption as per comments on Fig.1.

Period at end of caption (same in Fig. 4)

## Page 25.

**Fig 4.** Same comments as for Fig 3.

## Page 26.

**Line 2.** Please use present tense

**Line 3.** "...Tevatron, corresponding..."

**Line 4.** "W bosons" (no hyphen)

An 8 lines conclusion to a PRD on such an important analysis is highly inappropriate.

Please expand with quantitative comments on: (1) improvements of this analysis over the last CDF one, (2) consistency with past CDF results, (3) consistency with past and current other experiments' results, (4) what are the peculiarities of this work and why it's worth getting it published in CDF while much better results are available at LHC, (5) discuss comparison with theory.

You should also say somewhere in the conclusions that this measurements supersedes the one of Ref. [10]

## Page 27.

**Title:** remove one "Appendix", and "Cross section tables." (no capitals).

**Line 1.** Some short introductory text may be added to introduce what's presented in this appendix.

**Tables IX and X, captions.**



Remove “respective”

“...total systematic uncertainty which does not include the...”

“...jet multiplicity cross section for events in which  $|\eta_e| < 1.1$ ...”

c

“...for 0--4 inclusive number of jets.”

“...the uncertainty includes three contributions, the statistical component, the total systematic uncertainty, and the uncertainty associated with the knowledge of the integrated luminosity.”

On 3<sup>rd</sup> line of all tables from IX to XX, replace “in to” with “into”. On the 4<sup>th</sup> line replace “which is without” with “excluding”.

## Page 28.

**Tables XI and XII.** “...is divided into statistical uncertainty (central column) and systematic uncertainty (last column).”

“... $W$  cross section is restricted to events in which  $|\eta| < \dots$ ”

Delete “respectively” at the end.

## Pages 29, 30, 31, 32, 33, 34

**Tables XIII, XIV, XV, XVI, XVII, XVIII, XIX, XX captions.**

“...total systematic uncertainty which does not include the...”

“ $W$  cross section is restricted to events in which  $|\eta| < \dots$ ”,

“...the uncertainty includes three contributions, the statistical component, the total systematic uncertainty, and the uncertainty associated with the knowledge of the integrated luminosity.”

Also, remove “itself” at the end.

## Page 35.

**Line 1.** “J. Campbell and R.K. Ellis”.

**Line 2.** “R.K. Ellis, K. Melnikov, and G. Zanderighi.” (twice).

**Line 3.** Use standard AIP format for JHEP references (see SPRG guide).

**Line 11.** Use standard abbreviations and mention only first page.

**Line 15.** Use standard AIP format for JHEP references (see SPRG guide).

**Line 16.** If this paper has  $\leq 10$  authors please list them all.

**Line 17.** Please use FERMILAB-THESIS number if existing or get one if not.

**Line 10.** “D0 Collaboration”.

**Line 22.** Use standard SPRG format for paper abbreviation.

**Line 24.** There are extra digits in the volume number.

**Line 28.** Missing space between A and 268.

**Line 29.** Missing space between A and 494

## Page 36.

**Line 1.** No comma after “A”. Put a semicolon, not a period before K. Anikeev. Also, there's something wrong in this Computer Physics Communications paper since the authors don't include K. Anikeev

**Line 9.** Use standard format with collaboration between parentheses.

**Line 13.** Put the formula “ $M_T = \dots$ ” between commas.

**Line 17.** Delete hep-ex reference for published articles.

**Line 18.** Use standard format with collaboration between parentheses